Abstract: The intention of this research was to automate the process of formative assessment of e-learners using concept map (CM). Formative assessment assesses the learners during the learning process and further modifies the current teaching process to improve the learner’s outcome. It is specifically aimed at generating a performance feedback to improvise and accelerate learning. Formative assessment of the learners on the topics being taught is essential in an e-learning environment. Based on the learner’s learning, the e-learning system can change / update the pedagogy, recommendations can be made for further study, and the learners’ performance can be evaluated. In this paper, we propose CM based formative assessment of learners' learning. CM is an effective tool for determining what a learner knows in the topics covered. The CM constructed by the learner is mapped with the reference CM created by the subject expert. The experimental evaluation is presented and discussed.

Key words: Concept map (CM), e-assessment, e-learning, formative assessment.

1. Introduction

The teaching-learning process is the heart of any learning system. It is the most powerful instrument that helps in achieving the aims and objectives of education. Teaching and learning are related terms. In the teaching-learning process, the teacher, the learner, the curriculum and the pedagogy are organized systematically to attain some predetermined goal [1]. Learning refers to what the learner understands from the subject being taught. Teaching is providing proper direction and management of learning path. It is the process of providing opportunities for learners to learn. Assessment lies at the heart of the learning experience: how the assessment shapes their understanding of the curriculum and determines their ability to progress [2]. It is an essential element in teaching, learning processes. It is therefore not surprising that almost all learning management systems (LMSs) offer support for assessment, e.g., For the creation, execution, and evaluation of multiple choice tests. The term assessment is often used to summarize all activities that teachers use to help learners learn and to quantify the learning progress and outcomes. The latter, in particular, means that assessment measures and documents the knowledge, skills, and attitudes of an individual learner, a learning community (e.g., class, course, or workshop), or an educational institution. In the traditional system, there will be face-to-face communication between learners and the teacher. The teacher not only teaches, but also monitors and assesses the understanding of all the learners.

In any e-learning system, it is important that both teacher and learner keep track of learner’s progress and to check whether the learner has understood the topic. Hence, it would be more accurate and
comfortable if the assessment process is automated. One of the main advantages of e-learning is that it can facilitate adaptive learning such that instructors can dynamically revise and deliver instructional materials in accordance with the learners’ current progress. In general, adaptive teaching and learning refers to the use of what is known about learners, a priori or through interactions, to alter how a learning experience unfolds, with the aim of improving learners’ success and satisfaction [3], [4]. In e-learning, a regular knowledge assessment is required to be carried out using different kinds of tests for many reasons. Effective assessment and feedback can be considered as a practice that equips learners to study and perform to their best advantage in the complex disciplinary fields of their choice, and to progress with confidence and skill as lifelong learners. There are two types of assessment: Formative Assessment and Summative Assessment [5]. The goal of formative assessment is to monitor learner learning on a topic and the result can be used for both teacher and learner. We can provide ongoing feedback to the learners and the learners identify their strengths and weaknesses. It helps the instructors to update the pedagogy. Overall the formative assessment can improve the teaching learning process and the learners will get a better learning experience [6]. The goal of summative assessment is to evaluate learner learning at the end of an instructional unit by comparing it against some standard or benchmark. Result from summative assessments can be used formatively when learners or faculty use it to guide their efforts and activities in subsequent courses.

1.1 Bloom’s Taxonomy

Bloom's taxonomy is a classification of learning objectives within teaching learning process proposed in 1956 by a committee of educators chaired by Benjamin Bloom. It refers to a classification of the different objectives that teachers set for learners (learning objectives) [7]. Bloom's taxonomy divides teaching-learning objectives into three domains: Cognitive, Affective, and Psychomotor (sometimes loosely described as knowing/head, feeling/heart and doing/hands respectively). Within the domains, learning at the higher levels is dependent on having attained prerequisite knowledge and skills at lower levels. The goal of Bloom’s taxonomy is to motivate teachers to focus on all three domains, creating a more holistic form of education. Fig. 1 shows the levels in Bloom's taxonomy.

The aim of summative assessment is to evaluate learning at the lower levels of Bloom’s Taxonomy. But, formative assessment is to evaluate learning at the higher levels of Bloom’s Taxonomy. It can be understood that in formative assessment, one can evaluate in-depth learning of learners. It is also very important to assess learners in e-learning at the higher levels of Bloom’s taxonomy [8].

1.2 CM as a Knowledge Assessment Tool

CM (CM) is a pedagogical tool developed by Novak in 1970s [9]. CMs are based on two cognitive
theories: Ausubel's assimilation theory [10] considering the conceptual nature of human learning and hierarchical organization of knowledge and Deese's associationist memory theory [11] advocating networked arrangement of concepts. A CM is a semi-formal knowledge representation tool visualized by a graph consisting of finite, non-empty set of nodes, which depict concepts, and finite, non-empty set of arcs (directed or undirected), representing the relationship between concepts. A linking phrase can specify the kind of a relationship between concepts. As a rule, natural language is used to represent concepts and linking phrases. Moreover, all arcs of the graph can have the same weight, or weights can be different. A proposition, concept-relationship-concept triple -- is a semantic unit of CMs. It is a meaningful statement about some object or event in a problem domain [12]. According to Novak and Cañas [9], CMs are represented in a hierarchical fashion with the most general concepts at the top of the map and the more specific concepts below them, but cross-links can be used to indicate relationships between concepts in different domains of a CM by forming some kind of a network. Concept map has been successfully employed in many researches to help instructors and learners organize relationships among concepts. There are a wide variety of CM tasks which allow providing of knowledge assessment, adapted to the learners' characteristics. However, two main groups of them are:

- 'fill-in-the-map' tasks where the structure of a CM is given, and a learner must fill it using the provided set of concepts and/or linking phrases
- 'construct-the-map' tasks where a learner must decide on the structure of a CM and its content by him/herself.

Assessments based on "construct-the-map" tasks more accurately evaluate differences in the learners' knowledge structures and elicit higher-order cognitive processes [13]. "Construct-the-map" tasks are better than "fill-in the blank" tasks for capturing learners' partial knowledge. Based on their characteristics, if used as an assessment tool, "construct-the-map" tasks are more suitable for formative assessment, while "fill-in the blank" tasks are a better fit for large scale assessment [14]. Hence, in this research "construct-the-map" task is being used.

CMs allow evaluation of higher order levels of cognitive development in Bloom's taxonomy, especially when learners must choose the most prominent and most useful linking phrases and cross-links [15].

2. Related Work

Assessment is very critical in the teaching-learning process of any learning system to make sure that the learners are progressing towards the course objectives and to adjust the teaching-learning process accordingly. It does not only assign grades, but also, improve the quality of instruction in the teaching-learning process [16]. There are two types of assessment summative and formative assessment. Summative assessment is a periodic assessment used to determine at a particular point in time what learners' know and do not know. Formative assessment is part of the teaching-learning process that provides information needed to adjust teaching and learning while they are happening [17]. Classroom assessment techniques (CATs) are teaching strategies that provide formative assessments of learner learning. It has been argued that the use of CATs enhances and improves learner learning [18]. Formative assessment can have a powerful impact that motivates learners and leads to achievement [19]. Cauley and McMillan [20] discusses the key practices that teachers can use to gather important information about learner's understanding, provide feedback to learners, and enable learners to set and attain meaningful learning goals.

CMs are excellent tools to provide instructors with diagnostic pre-assessment prior to beginning a unit and formative assessments during learning activities [21]. Wu [22] indicate that CMs develop learner abilities in certain critical areas like, the ability to draw reasonable inferences from observations, the ability to synthesize and integrate information and ideas, the ability to learn concepts and theories in the subject
area. They were chosen as a strategy to empower learners to be more effective readers and knowledge creators [22]. CM can also be used to enhance the interaction of teaching and learning with the goal to foster higher order thinking, i.e., analyzing a problem situation, evaluating possible solutions, and creating innovative ideas for problem solving. CM based teaching and learning can improve learner participation in higher order thinking activities [23].

2.1. Existing Systems for Automatic Assessment

Table 1 illustrates some existing systems available for automatic assessment of learners.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Existing System for automatic assessment</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Scheme-robo [25]</td>
<td>Assessment of programming assignments in scheme</td>
</tr>
<tr>
<td>3</td>
<td>AutoGrader [26]</td>
<td>Assessment of Java programs,</td>
</tr>
<tr>
<td>5</td>
<td>MyCodeMate [<a href="http://www.mycodemate.com">http://www.mycodemate.com</a>]</td>
<td>Support the assessment of several programming languages</td>
</tr>
<tr>
<td>6</td>
<td>Course Marker [27]</td>
<td>Summative assessment</td>
</tr>
<tr>
<td>7</td>
<td>BOSS [28]</td>
<td>Summative assessment</td>
</tr>
<tr>
<td>8</td>
<td>AT(x) [29]</td>
<td>Summative assessment</td>
</tr>
<tr>
<td>9</td>
<td>Moodle [<a href="http://moodle.org">http://moodle.org</a>]</td>
<td>Questionnaires, Multiple-choice tests, File uploads</td>
</tr>
<tr>
<td>10</td>
<td>Blackboard [<a href="http://www.blackboard.com/">http://www.blackboard.com/</a>]</td>
<td>Questionnaires, Multiple-choice tests, File uploads</td>
</tr>
<tr>
<td>11</td>
<td>OLAT [<a href="http://olat.org">http://olat.org</a>]</td>
<td>Questionnaires, Multiple-choice tests, File uploads</td>
</tr>
</tbody>
</table>

Almost all of these systems have one common property of providing (along with the actual testing of programming assignments) functionality for managing users, courses, assignments, and submissions. Furthermore, these systems are difficult to extend and adapt to one's own requirements. They are built for the purpose of testing programs in a certain language or employ a certain test method. This results in rather inflexible, monolithic systems that cannot accommodate functional extensions. The end result of the assessment provided by these systems can be used only for grading the learners. All these systems perform only summative assessment but not formative assessment. Most of these systems are used to assess only the programming exercises which cannot be extended for other contexts.

3. Proposed Model

This paper proposes a model for the automatic formative assessment of learner learning by comparing the CMs developed by the learner and created by the expert. There are four models in any e-learning system, i) Teacher Model ii) Learner Model iii) Pedagogical Model iv) Score Engine (Assessment Model). The process begins with the Teacher Model, through which the subject expert (teacher) generates the course content and also the “reference CM” for the content learnt by the learner.

Learner Model provides a necessary user interface for interaction and presents the learning content to the learner. After learning the subject, the learner generates the CM on the subject learned. This CM is given as an input to the Score Engine, which is the prime focus of this paper. Score Engine maps the reference CM with the CM generated by the learner and returns the score for learner learning. This learning assessment can be used by Learner Model to learn the learner and in turn Pedagogical Model can decide the learning
path. Fig. 2 gives the architecture of the proposed model. The system is implemented as a Web-based three-tier client-server application consisting of the following architectural layers:

- A data storage layer represented by database management system PostgreSQL
- An application logics layer composed of the application server (Apache Tomcat), server-side code, and a special persistence and query framework – hibernate
- The representation layer based on Swing, JGraph, and JGoodies libraries.

### 3.1. Preprocessing

The CM generated by both teacher and learner is preprocessed to achieve a better performance when using Natural Language Processing (NLP) techniques to find the similarity between two concepts. The preprocessing is aimed at filtering the input text in each concept and their relationship. The tokenization is performed using the OpenNLP parser. In this process, tokens are extracted using regular expressions to recognize whitespaces, punctuation marks, hyphens, and numbers, among others. Tokenization sometimes faces challenges in finding the starting and end point of a token. This is because not all tokens are made of one word. For example, 'operating systems' must be chunked to be a single word and not two tokens. Another example, 'Java Programming' must be chunked to be a single word and not two tokens. We employed the GATE tool tagger which is rule based to predict the part-of-speech (POS) of each word. Some of the transformation rules used for tagging are:

- Converting a noun to a number (CD) if '.' appears in the word
- Converting a noun to a past participle if ((string)words[i]) ends with 'ed'
- Converting any type to adverb if it ends in 'ly'.

We used the following three main techniques for preprocessing:

- Stop words removal that is dropping all prepositions, conjunctions, and other parts of speech that are very common in the texts
- Stemming that is keeping just the root of each word to allow statistical techniques devising the same meaning from different inflections.
• POSTagging is to assign a part-of-speech to each word in a sentence. A unique tag to each word reduces the number of parses.

3.2. Score Engine

The score engine takes the pre-processed CM constructed by the learner and the CM generated by the expert as inputs. The score engine calculates and returns the score value of the learning of the learner. The score value is a result of comparison of the CM constructed by the learner and the CM generated by the expert. CM’s can be viewed as a directed graph with weighted links where, a graph is a representation of a set of concepts where some pairs of concepts are connected by links. The weight of the link denotes the importance of the relationship between the two concepts. For instance, consider the CM given in Fig. 4 for Java Language, the first level concepts consists features of the Java language. The expert may give more importance to the feature OOP than the rest, i.e. the link connecting the concepts Java language and OOP will have more weight than the links connecting Java and other features. The score should be calculated based on the weight between pairs of concepts if it exists in the CM constructed by the learner. This comparison and score updation should be done at each level. A level consists of the set of concepts that are generated by another concept. Let Y be the CM generated by the learner and X be the CM generated by the expert. The algorithm for calculating the score is given in Fig. 3.

```
Calc_Score(X, Y)
Step 1:   Create_Queue Qx and Qy
Step 2: Initialize Score with 0
Step 3: If (Sim(Root(X),Root(Y)) != 0)
       Qx.Enqueue(Root(X))
       Qy.Enqueue(Root(Y))
End If
Step 4: Nx = Qx.Deque()
       Ny = Qy.Deque()
Step 5: While (Successor(Nx) != NULL)
       Si(Nx) is the i-th successor of Nx;
       Initialize Max = 0
       While (Successor(Ny) != NULL)
       Sj(Ny) is the j-th successor of Ny
       If (Sim(Si(Nx), Sj(Ny)) > Max) then
       Max = Sim(Si(Nx), Sj(Ny))
       Pos = j
       End if
       End While
       If (Max > 0) then
       Qx.Enqueue(Si(Nx))
       Qy.Enqueue(Sj(Ny))
       Score += Wt(R(Nx, Si(Nx))) * ((Sim(Nx, Ny) + Max) / 2)
       End If
       End While
Step 5: If (IsEmpty(Qx) OR IsEmpty(Qy))
       Return score
Else
       Repeat From Step 4
End if
```

Fig. 3. Algorithm for calculating score.

The algorithm is based upon breadth first search (BFS) traversal, a strategy for searching in a graph. The BFS begins at a root (begins from the concept Java in Fig. 4) and inspects all the neighboring nodes (all the features of Java in the Fig. 4). Then for each of those neighbor nodes in turn, it inspects their neighbor
nodes, which were unvisited, and so on. BFS is applied to both the graphs X and Y simultaneously based on the similarity of the contents of the node. In this proposed algorithm we maintain two queues, one corresponding to X and the other corresponding to Y. At each level if the node similarity is greater than 0, then the child nodes are inserted in the tree, at the same time dequeuing the current node and incrementing the score of Y with respect to X based on the similarity. The same process continues for all the nodes till any of the queue becomes empty. The algorithm also takes into account that any node in X can correspond to a maximum of one node in Y i.e. there cannot be two nodes describing the same concept. The algorithm selects the one with maximum similarity with corresponding node in Y.

3.3. Feedback and the Purpose of the Numeric Score

It is widely recognized that feedback is an important part of the formative assessment. Effective feedback helps learners to develop their understanding and improve their performance in relation to the goals of learning. After comparison of the CM constructed by the learner and the CM generated by the expert, the system will provide feedback to learners with information about the strengths and weaknesses of responses, the outcomes achieved and learners’ performance in relation to the goal of learning and to other learners. The system will automatically annotate the learners CM. The system will also provide some meaningful information about the CM constructed by the learners and also correction of misunderstandings.

In e-learning environment, it is not only enough to give feedback to the learners, but also necessary to make changes in pedagogy to provide an adaptive learning environment. The numeric score calculated is used by the learner model to predict the performance of the learners and pedagogical model to make suitable changes in the pedagogy and to make suitable recommendations accordingly. The numeric score can also be used to compare the performance with other learners.

4. Empirical Results and Discussion

Consider the CM generated for Java language [30] given in Fig. 4.
The CM given in Fig. 4 created by a subject expert consists of 33 concepts and 33 links with Java Language as the root. Let us consider that each Link has a weight of 10. Total weight therefore is 330. In case the learner has missed the feature ‘secure’ in the CM generated, the algorithm will not add the weight of the ‘secure’ node as well as any of its children nodes. Hence the weight of Y with respect to X will be \((280/330)\times100\) assuming the similarity of the existing nodes is 1. The score engine will return 84.7% as the score of the learner. This indicates that the knowledge level of the learner in the learned topic is 84.7%. This can be used by the learner model to predict the performance of the learners and pedagogical model can make suitable changes in the pedagogy and can make suitable recommendation.

In this study, a CM approach is proposed for formative assessment of e-learners. Formative assessment supports and improves teaching learning process in e-learning. Based on the score value, the e-learning system can change the pedagogy and recommend learning path to the learner to get a better learning experience. To evaluate the effectiveness of the approach, two experiments were conducted. First experiment was conducted to find the effectiveness of CM as an assessment tool. It is a questionnaire based experiment. A total of 73 learners doing “Java Programming” Course participated in this study. Five-score Likert-type scale items, which ranged from “Strongly Disagree” to “Strongly Agree” was used for analyzing the level of satisfaction. All learners learned the topic features of Java. After completing learning, the learners were asked to generate CM for the topic learned by them. Out of 73, 51 learners supported CM as learning tool. The prime reason given by them is that this learning approach can help learn the contents from a new perspective and this learning system enables better understanding of the learning content. They are happy with the feedback and explanation of each concept given by the system. They are able to understand the concept more clearly. All learners strongly believe that the online feedback given as part of formative assessment helps them to focus on and achieve their learning goals in the field. We have analyzed the unexpected results as well. The main challenge faced by the learners is that this approach requires thinking differently about the learning content and this makes the current learning activity more challenging. They are not able to link the concepts.

Table 2 summarizes the survey results and gives a more detailed statistical study (with the mean and the standard deviation) of the different items in the study. In summary, the results indicate that the learners’ level of satisfaction and learning experience in e-learning environment that uses formative assessment was ≈4.3. In general terms, the survey data shows the learning experience was evaluated positively by learners.

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction Level</td>
<td>4.301</td>
<td>0.725</td>
</tr>
<tr>
<td>Overall learning experience</td>
<td>4.205</td>
<td>0.517</td>
</tr>
<tr>
<td>Importance of Formative assessment in e-learning environment</td>
<td>4.524</td>
<td>0.779</td>
</tr>
<tr>
<td>Usefulness of the feedback given</td>
<td>4.728</td>
<td>0.518</td>
</tr>
<tr>
<td>Effectiveness of the personalized learning environment after formative assessment</td>
<td>3.75</td>
<td>0.952</td>
</tr>
<tr>
<td>Improvement in academic performance</td>
<td>4.103</td>
<td>0.871</td>
</tr>
<tr>
<td>Enabled to attain meaningful learning goal</td>
<td>4.271</td>
<td>0.729</td>
</tr>
<tr>
<td>Confidence in subject</td>
<td>4.132</td>
<td>0.785</td>
</tr>
</tbody>
</table>

Second experiment was designed to evaluate the effectiveness of our approach. However, “fill-in the blank” based CMs can be scored more efficiently than “construct-the-map” based CMs. There is no need to evaluate the effectiveness of the proposed approach if fill-in the blank based CMs is used. Because in the fill-in the blank based CM, the concepts and interlinks are predetermined. But this research uses construct-the-map based CMs. In this the concepts and interlinks may vary. Hence it is necessary to evaluate the
effectiveness of the system. Two standard measures from information retrieval namely precision and recall are used to evaluate our approach. They are formulated as below:

Precision: it is the probability that if a random Learner CM Li is scored Si, this decision is correct. It can be viewed as the “degree of soundness” of the system. That is

\[
\text{Precision} = \frac{A}{A + B}
\]

Recall: it is the probability that if a random Learner CM Li to be scored under Si, this decision is taken. It can be viewed as the degree of completeness of the system. That is

\[
\text{Recall} = \frac{A}{A + C}
\]

The contingency table consists mainly of the following values

- A: The number of learner CMs the system correctly assesses with a particular score value (true positives),
- B: The number of learner CMs the system incorrectly assesses with a particular score value (false positives),
- C: The number of learner CMs that belong to the particular score value but which the system does not assign that score value (false negatives)
- D: The number of learner CMs the system correctly does not assign score value (true negative),

Precision-Recall (PR) curves give a better representation of an algorithm's performance. The precision-recall curve is given in Fig. 5 and the values are tabulated in Table 3. For each learner, the precision values are computed at 11 standard recall levels, 0%, 10%, ..., and 100%. Finally, the precision-recall curve is plotted. The computation of these measures depends essentially on a contingency table obtained from the results of the formative assessment of all the learners.

| Precision | 0.94 | 0.912 | 0.87 | 0.837 | 0.819 | 0.79 | 0.75 | 0.7356 | 0.722 | 0.719 | 0.712 |
| Recall    | 0    | 0.1   | 0.2  | 0.3   | 0.4   | 0.5  | 0.6  | 0.7   | 0.8   | 0.9   | 1     |

Fig. 5. Precision – Recall Curve for the proposed model.

From Fig. 5 we observed that better precision is obtained when recall values are < 0.2 and the precision values are moderately good when the recall values lies between 0.2 and 0.6. However the precision values are still greater than 0.7 when the recall values approaches to 1. This shows that in our proposed system, on average, a better precision value is always obtained.
5. Conclusion

The system discussed in this paper is an ideal environment for automated formative assessment of e-learners. This paper focuses on how CM can be used for formative assessment of learners in e-learning system. The effectiveness of teaching learning process in the e-learning system depends on appropriate teaching and knowledge assessment methods. CM provides valuable information for teaching learning process. Experimental results show that our proposed model can assess the learning of learners effectively. Moreover, the results of this study indicate that the use of the formative assessment in e-learning has important effects on the learners’ learning experience and academic outcomes. A hidden advantage that was evident from the learners’ responses was that this could be used for self-assessment of learners also. This enabled the learners to attain meaningful learning goal. In e-learning environment, it is also necessary to see that all learners are motivated and self-worth when realised in a context of completing the course. It will be more effective if we automatically generate the reference CM for the chosen topic.

References

V. Senthil Kumaran was born in 1974 in India. He received his master degree in computer science from Madurai Kamaraj University, Madurai, India and Ph.D in computer applications from Anna University, Chennai, India. He is currently working as a senior assistant professor at the Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Coimbatore, India. His current research interests include e-learning, knowledge management, semantic web, data mining, and intelligent information retrieval.